Prognostics Framework

Abhinav Saxena, *Member, IEEE,* Jose Celaya, *Member, IEEE,*
Kai Goebel, Bhaskar Saha, *Member, IEEE,* Sankalita Saha, *Member, IEEE*

I. Prognostics Terms and Definitions

In this section we describe some commonly used terms in prognostics. Similar terms have been used interchangeably by different researchers and in some cases the same term has been used to represent different notions. This list is provided to reduce ambiguities that may arise by such non-standardized use.

A. Assumptions

- Here prognostics is considered to be the detection of failure precursors and the prediction of RUL based on the current state assessment and expected future operational conditions of the system.
- It is possible to estimate a health index as an aggregate of features and conditions
- RUL estimation is a prediction/ forecasting/ extrapolation process.
- Algorithms under consideration are capable of generating a single RUL value for each prediction. That is, algorithms that produce RUL distributions can be adapted to compress the distribution to a single estimated number for comparison purposes.
- All systems are under continuous monitoring and have the measurement capability that can acquire data as a fault evolves.

B. Glossary

RUL: Remaining Useful Life – amount of time left before system health falls below a defined failure threshold

UUT: Unit Under Test

i : Index for time instant t_i

EOL: End-of-Life - Time index of actual end of life

EOP: End-of-Prediction – earliest time index, i, when prediction has crossed the failure threshold

0: Time index for time of the birth of the system, t_0

F: Time index for the time when fault occurs, t_F

D: Time index at which the fault is detected by diagnostic system, t_D

P: Time index at which the first prediction is made by the prognostic system, t_P

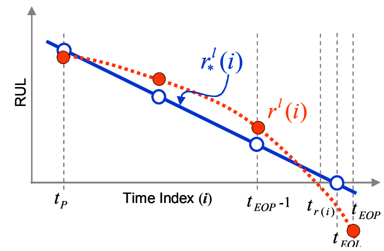

Figure 1. Illustration depicting some important prognostic time definitions and prediction concepts.

$f_n^l(i)$: Value of the n^{th} Feature for the l^{th} UUT at time index i

$c_n^l(i)$: Value of the n^{th} operational condition for the l^{th} UUT at time index i

$r^l(i)$: RUL Estimation at time t_i given that data is available up to time t_i for the l^{th} UUT

$\pi^l(i\,|\,j)$: Prediction at time index i given data up to time t_j for the l^{th} UUT. Prediction may be made in any domain, e.g. feature, health, etc.

$\Pi^l(i)$: Trajectory of predictions at time index i for the l^{th} UUT

$h^l(i)$: Health of system for the l^{th} UUT

Definition 1 - ***Time Index:*** The time in a prognostics application can be discrete or continuous. We will use a time index i instead of the actual time, e.g., i=10 means t_{10}. This takes care of cases where sampling time is not uniform. Furthermore, time indexes are invariant to time-scales.

Definition 2 - ***Time of Detection of Fault:*** Let D be the time index (t_D) at which the diagnostic or fault detection algorithm detected the fault. This process will trigger the prognostics algorithm which should start making RUL predictions shortly after the fault was detected as soon as enough data has been collected. For some applications, there may not be an explicit declaration of fault detection, e.g., applications like battery health management, where prognosis is carried out on a decay process. For such applications t_D can be considered equal to t_0 (time of birth) i.e., we expect to trigger prognosis as soon as enough data has been collected and not wait for an explicit diagnostic flag (Figure 2).

Definition 3 - ***Time to Start Prediction:*** We will differentiate between the time when a fault is detected (t_D) and the time when the system starts predicting (t_P). For certain algorithms $t_D = t_P$ but in general $t_P \geq t_D$ as these algorithms need some time to tune with additional fault progression data before they can start making predictions (Figure 2). In cases where a data collection system is continuously collecting data even before fault detection, enough data is already available to start making predictions right away and hence $t_P = t_D$.

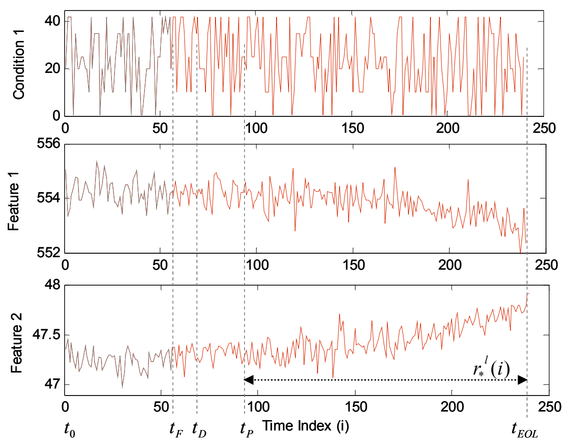

Figure 2. Features and conditions for l^{th} UUT.

Definition 4 - ***Prognostics Features:*** Let $f_n^l(i)$ be a feature at time index i, where $n = 1, 2, \dots, N$ is the feature number, and $l = 1, 2, \dots, L$ is the UUT index (an index identifying the different units under test). In prognostics, irrespective of the analysis domain, i.e., time, frequency, wavelet, etc., features take the form of time series and they can be physical variables, system parameters or any other quantity that can be computed from measurable variables of the system that provides or aides the prognosis. The features can be also referred to as a feature vector $F^l(i)$ of the l^{th} UUT at time index i.

Definition 5 - ***Operational Conditions:*** Let $c_m^l(i)$ be an operational condition at time index i, where $m = 1, 2, \dots, M$ is the condition number, and $l = 1, 2, \dots, L$ is the UUT index. The operational conditions describe how the system is being operated and are sometimes referred to as the load on the system. The conditions can also be referred to as a vector $C^l(i)$ of the l^{th} UUT at time index i.

Definition 6 - ***Health Index:*** Let $h^l(i)$ be a health index at time index i for UUT $l = 1, 2, \dots, L$. h can be considered a normalized aggregate of health indicators (relevant features) and operational conditions.

Definition 7 - ***Ground Truth:*** Ground truth, denoted by the subscript $_*$, represents our best belief of the true value of a system variable. In the feature domain $f_{*n}^l(i)$ may be directly or indirectly calculated from measurements. In the health domain, $h_*^l(i)$ is the computed health at time index i for UUT $l = 1, 2, \dots, L$ after a run to failure test. This health index represents an aggregate of information provided by features and operational conditions up to time index i.

Definition 8 – ***History Data:*** History data, denoted by the subscript $_\#$, encapsulates all the information we know about a system *a priori*. Such information may be of the form of archived measurements or EOL distributions, and can refer to variables in both the feature and health domains represented by $f_{\#n}^l(i)$ and $h_\#^l(i)$ respectively.

Definition 9 - ***Point Prediction:*** Let $\pi^l(i \mid j)$ be a point prediction of a variable of interest at time index i given information up to time t_j, where $t_j \leq t_i$. $\pi^l(i \mid j)$ for $i = EOL$ represents the critical threshold for a given health indicator. Predictions can be made in any domain, features or health. In some cases it is useful to extrapolate features and then aggregate them to compute health and in other cases features are aggregated to a health and then extrapolated to estimate RUL.

Definition 10 - ***Trajectory Prediction:*** Let $\Pi^l(i)$ be the trajectory of predictions at time index i such that $\underline{\Pi}^l(i)$ = $\left\{\pi^l(i \mid i), \pi^l(i+1 \mid i), \dots, \pi^l(EOL \mid i)\right\}$

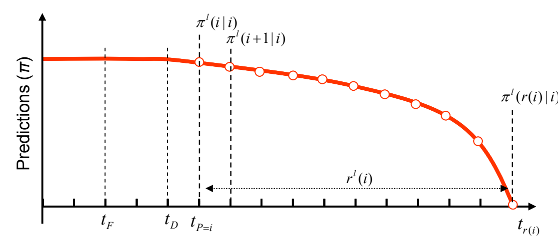

Figure 3. Illustration showing a trajectory prediction. Predictions may modify every time instant and hence the corresponding RUL estimate.

Definition 11 - ***RUL Estimation:*** Let $r^l(i)$ be the remaining useful life estimation at time index i given that the information (features and conditions) up to time index i and an expected operational profile for the future are available. As shown in Figure 4, prediction is made at time t_i and it predicts the RUL given information up to time i for the l^{th} UUT RUL will be estimated as $r^l(i) = \arg_z\left\{h(z) = 0\right\} - i$.

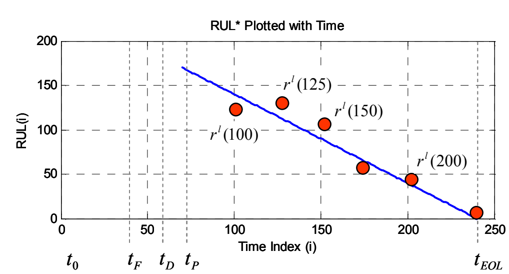

Figure 4. Comparing RUL predictions from ground truth (t_p [70,240], t_{EOL} = 240, t_{EOP} > 240).

II. FORECASTING APPLICATIONS CLASSIFICATION

Based on our survey of several forecasting application domains, we identified two major classes of forecasting applications (see Figure 5).

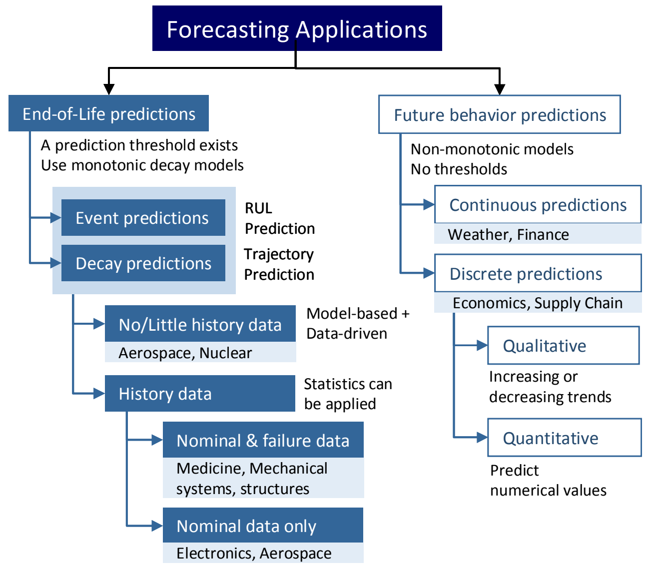

Figure 5. Different categories of the forecasting applications

In one class of applications a prediction is made on a continuous basis, and the trend of data is generally non-monotonic. These predictions may be discrete (e.g. forecasting market demand for a particular month) or continuous (e.g. variation of temperature over the period of next week). Predictions can be quantitative (e.g. prediction of exact numbers) or qualitative (e.g. high or low demands) in nature. Applications like weather and finance have been in existence for quite a while and have matured to a good extent. The other class of applications involves the existence of a critical threshold such that the system under test is declared to have lost a defined degree of functional capability (including complete failure) if it crosses the threshold. These applications usually can be modeled using decay models. Here the task of prognostics is to predict a RUL estimate. In some cases, where enough history data exists (e.g. medicine) or can be experimentally generated (e.g. mechanical systems) for nominal and failure conditions, a variety of data-driven or statistical techniques can be applied. In such situations it is also relatively easy to evaluate the performance by comparing the prediction *a posteriori*. However, there are critical applications where run-to-failure experiments cannot be afforded and very little failure history data is available (e.g. aerospace). In such cases a variety of methods based on data-driven and model-based techniques have been proposed. It becomes extremely tricky and difficult to assess the performance in such cases due to absence of knowledge about the future outcomes. Methods are tested on experimental or simulated data and are expected to perform on real systems. Unfortunately algorithm performance does not always translate meaningfully from one dataset to another or one domain to another. Therefore, a standard set of metrics independent of application domain would be very desirable.